

The Insulating State of Matter: A Geometrical Theory

Raffaele Resta

*Dipartimento di Fisica, Università di Trieste, Italy,
and DEMOCRITOS National Simulation Center,
IOM-CNR, Trieste, Italy*

(Submitted as a “Colloquium paper” to Eur. Phys. J. B)

In 1964 W. Kohn published the milestone paper “Theory of the insulating state”, according to which insulators and metals differ in their *ground state*. Even before the system is excited by any probe, a different organization of the electrons is present in the ground state and this is the key feature discriminating between insulators and metals. However, the theory of the insulating state remained somewhat incomplete until the late 1990s; this review addresses the recent developments. The many-body ground wavefunction of any insulator is characterized by means of geometrical concepts (Berry phase, connection, curvature, Chern number, quantum metric). Among them, it is the quantum metric which sharply characterizes the insulating state of matter. The theory deals on a common ground with several kinds of insulators: band insulators, Mott insulators, Anderson insulators, quantum Hall insulators, Chern and topological insulators

Contents

Introduction	1
Geometry in quantum mechanics	2
Phases and distances	2
Berry phase	3
Connection and curvature	3
Chern number	4
Metric	4
Sum over states	5
Many-electron wavefunction	5
Boundary conditions	5
Open boundary conditions	6
Periodic boundary conditions	6
The localization tensor	7
Discretized formulae	7
Sum over states again	8
Conductivity	8
Linear response	8
Kubo formula for conductivity	9
Sum rules	9
Screened vs. unscreened field	9
Localization in the insulating state	10
Independent electrons	11
Band insulators and band metals	11
Wannier functions	12
Chern insulators	12
Localization in different kinds of insulators	12
Small molecules	12
Band insulators	12
Correlated (Mott) insulators	13
Disordered (Anderson) insulators	14
Quantum Hall insulators	14
Chern insulators	16
Topological insulators	16
Conclusions	17
References	17

INTRODUCTION

The standard textbook approach to the insulating state of matter is based on band theory. Following Bloch’s theorem [1] in 1928, the main result is due to Wilson in 1931 [2]. The single-particle spectrum of a lattice-periodical Hamiltonian is in general gapped, and the electron count determines where the Fermi level lies. If it crosses a band one has a conductor: an applied electric field induces free acceleration of the electrons (at $T = 0$ in absence of dissipation). If the Fermi level lies instead in a gap, one has an insulator: in presence of a field the electronic system polarizes, but no steady-state current flows for $T \rightarrow 0$. This very successful theory explains the insulating/conducting behavior of most common materials across the periodic table, for which band structure calculations became soon available.

At the root of band theory are two basic assumptions: the electrons are noninteracting (in a mean-field sense), and the solid is crystalline. By the early 1960s, however, it became clear that there are solids where these two assumptions are very far from the truth, and where the insulating behaviour is due to completely different mechanisms. The works of Mott in 1949 [3] and of Anderson in 1958 [4] opened new avenues in condensed matter physics. In the materials which we now call Mott insulators the insulating behaviour is due to electron correlation [5], while in those called Anderson insulators it is due to lattice disorder [6].

In a milestone paper appeared in 1964 Kohn [7] provided a more comprehensive characterization of the insulating state of matter, which encompasses band insulators, Mott insulators, Anderson insulators, and eventually any kind of insulating material. According to Kohn, the electrons in the insulating state satisfy a many-electron localization condition [8]. This kind of localization must be defined in a subtle way given

that, for instance, the Hamiltonian eigenstates in a band insulator are obviously *not* localized. According to the original Kohn's formulation, the insulating behaviour arises whenever the ground-state wavefunction of an extended system breaks up into a sum of contributions which are localized in essentially disconnected regions of the many-electron configuration space.

Kohn's theory remained little visited for many years [9] until the 1990s, when a breakthrough occurred in electronic structure theory: the modern theory of polarization (for historical presentations, see e.g. Refs. [10–12]). Inspired by the fact that electrical polarization discriminates *qualitatively* between insulators and metals, Resta and Sorella [13] in 1999 provided a definition of many-electron localization rather different from Kohn's, and deeply rooted in the theory of polarization. Their program was completed soon after by Souza, Wilkens and Martin [14] (hereafter quoted as SWM), thus providing the foundations of the modern theory of the insulating state. An early review paper appeared in 2002 [15].

Several advances occurred afterwards, and the present review illustrates most of them. But the major aim of this work is to provide a thorough formulation of the modern theory of the insulating state by means of geometrical concepts, which remained somewhat hidden and implicit in the original literature. While the theory of polarization is based on a Berry phase, the theory of the insulating state as presented here is rooted in quantum distance and metric; most of the results previous published in different form will be revisited emphasizing their geometrical content.

GEOMETRY IN QUANTUM MECHANICS

Let us assume that a generic time-independent quantum Hamiltonian has a parametric dependence. The Schrödinger equation is

$$H(\boldsymbol{\kappa})|\Psi(\boldsymbol{\kappa})\rangle = E(\boldsymbol{\kappa})|\Psi(\boldsymbol{\kappa})\rangle, \quad (1)$$

where the d -dimensional real parameter $\boldsymbol{\kappa}$ is defined in a suitable domain of \mathbb{R}^d : a 2d $\boldsymbol{\kappa}$ has been chosen for display in Fig. 1. In this Section we start discussing the most general case, and therefore for the time being we do not specify which quantum system is described by this Hamiltonian, nor what the physical meaning of the parameter $\boldsymbol{\kappa}$ is. Later on, $|\Psi(\boldsymbol{\kappa})\rangle$ will be identified with a many-electron wavefunction. The state vectors $|\Psi(\boldsymbol{\kappa})\rangle$ are all supposed to be normalized and to reside in the same Hilbert space: this amounts to saying that the wavefunctions are supposed to obey $\boldsymbol{\kappa}$ -independent boundary conditions. We focus on the ground state $|\Psi_0(\boldsymbol{\kappa})\rangle$, and we assume it to be nondegenerate for $\boldsymbol{\kappa}$ in some domain of \mathbb{R}^d .

Any quantum mechanical state vector is arbitrary by a constant phase factor. Here we refer to choosing this

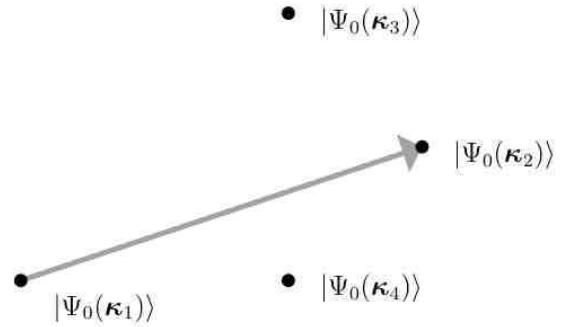


FIG. 1: State vectors in the two-dimensional $\boldsymbol{\kappa}$ -space. The phase difference between two of them is defined as $e^{-i\Delta\varphi_{12}} = \frac{\langle\Psi_0(\boldsymbol{\kappa}_1)|\Psi_0(\boldsymbol{\kappa}_2)\rangle}{|\langle\Psi_0(\boldsymbol{\kappa}_1)|\Psi_0(\boldsymbol{\kappa}_2)\rangle|}$, and their distance as $D_{12}^2 = 1 - |\langle\Psi_0(\boldsymbol{\kappa}_1)|\Psi_0(\boldsymbol{\kappa}_2)\rangle|^2$.

phase as to the choice of the gauge. All measurable quantities (e.g. expectation values) are obviously gauge-invariant, but the reverse is also true: all gauge-invariant properties are—at least in principle—measurable.

It is expedient to define the ground-state projector (a.k.a. density matrix) and its complement, i.e.

$$\hat{P}(\boldsymbol{\kappa}) = |\Psi_0(\boldsymbol{\kappa})\rangle\langle\Psi_0(\boldsymbol{\kappa})|; \quad \hat{Q}(\boldsymbol{\kappa}) = \hat{1} - \hat{P}(\boldsymbol{\kappa}). \quad (2)$$

Both $\hat{P}(\boldsymbol{\kappa})$ and $\hat{Q}(\boldsymbol{\kappa})$ are obviously gauge-invariant.

Phases and distances

We define the *phase difference* between the ground eigenstates at two different $\boldsymbol{\kappa}$ points in the most natural way:

$$e^{-i\Delta\varphi_{12}} = \frac{\langle\Psi_0(\boldsymbol{\kappa}_1)|\Psi_0(\boldsymbol{\kappa}_2)\rangle}{|\langle\Psi_0(\boldsymbol{\kappa}_1)|\Psi_0(\boldsymbol{\kappa}_2)\rangle|}; \quad (3)$$

$$\Delta\varphi_{12} = -\text{Im log } \langle\Psi_0(\boldsymbol{\kappa}_1)|\Psi_0(\boldsymbol{\kappa}_2)\rangle. \quad (4)$$

For any given choice of the two states, Eqs. (3) and (4) provide a $\Delta\varphi_{12}$ which is unique modulo 2π , except in the very special case that the states are orthogonal. However, it is also clear that such $\Delta\varphi_{12}$ is gauge-dependent and *cannot* have, by itself, any physical meaning.

The distance between quantum states has been defined by Bures [16] as:

$$D_{12}^2 = 1 - |\langle\Psi_0(\boldsymbol{\kappa}_1)|\Psi_0(\boldsymbol{\kappa}_2)\rangle|^2. \quad (5)$$

Such distance fulfills the familiar axioms from calculus textbooks; it vanishes when the two states physically coincide (i.e. independently of the phase factor), and is maximum (equal to one) when the states are orthogonal. At variance with $\Delta\varphi_{12}$ defined above, the Bures distance

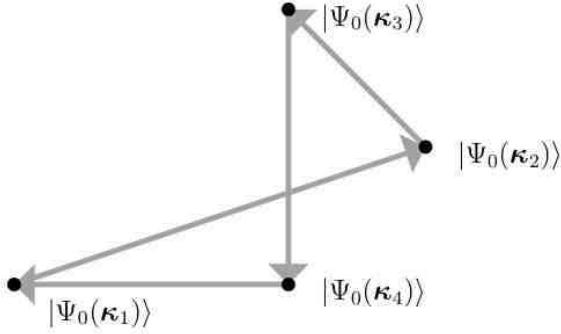


FIG. 2: A closed path joining four states in κ -space.

is gauge-invariant, and can be explicitly expressed in terms of ground-state projectors

$$D_{12}^2 = 1 - \text{Tr} \{ \hat{P}(\kappa_1) \hat{P}(\kappa_2) \}, \quad (6)$$

where “Tr” is the trace over the Hilbert space.

Berry phase

We have already observed that the phase difference $\Delta\varphi_{12}$ between any two states is gauge-dependent and cannot have any physical meaning. Matters are quite different when we consider the *total* phase difference along a closed path which joins several points in a given order, as shown in Fig. 2:

$$\begin{aligned} \gamma &= \Delta\varphi_{12} + \Delta\varphi_{23} + \Delta\varphi_{34} + \Delta\varphi_{41} \\ &= -\text{Im} \log \langle \Psi_0(\kappa_1) | \Psi_0(\kappa_2) \rangle \langle \Psi_0(\kappa_2) | \Psi_0(\kappa_3) \rangle \times \\ &\quad \times \langle \Psi_0(\kappa_3) | \Psi_0(\kappa_4) \rangle \langle \Psi_0(\kappa_4) | \Psi_0(\kappa_1) \rangle. \end{aligned} \quad (7)$$

It is now clear that all the gauge-arbitrary phases cancel in pairs, such as to make the overall phase γ a gauge-invariant quantity. The above simple-minded algebra leads to a result of overwhelming physical importance: in fact, a gauge-invariant quantity is potentially a physical observable. In essence, this is the revolutionary message of Berry’s celebrated paper, appeared in 1984 [17, 18].

Next we consider a smooth closed curve C in the parameter domain, such as in Fig. 3, and we discretize it with a set of points on it. Using Eq. (3), we write the phase difference between any two contiguous points as

$$e^{-i\Delta\varphi} = \frac{\langle \Psi_0(\kappa) | \Psi_0(\kappa + \Delta\kappa) \rangle}{|\langle \Psi_0(\kappa) | \Psi_0(\kappa + \Delta\kappa) \rangle|}. \quad (8)$$

If we further assume that the gauge is so chosen that the phase varies in a *differentiable* way along the path, then from Eq. (8) we get to leading order in $\Delta\kappa$:

$$-i\Delta\varphi \simeq \langle \Psi_0(\kappa) | \nabla_\kappa \Psi_0(\kappa) \rangle \cdot \Delta\kappa. \quad (9)$$

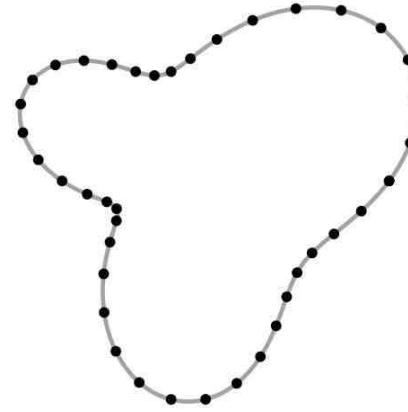


FIG. 3: A smooth closed curve C in κ -space, and its discretization.

In the limiting case of a set of points which becomes dense on the continuous path, the total phase difference γ converges to a circuit integral:

$$\gamma = \sum_{s=1}^M \Delta\varphi_{s,s+1} \longrightarrow \oint_C \mathbf{A}(\kappa) \cdot d\kappa, \quad (10)$$

where $\mathbf{A}(\kappa)$ is called the Berry connection:

$$\mathbf{A}(\kappa) = i \langle \Psi_0(\kappa) | \nabla_\kappa \Psi_0(\kappa) \rangle. \quad (11)$$

Since the state vectors are assumed to be normalized at any κ , the connection is *real*; we can therefore equivalently write

$$\mathbf{A}(\kappa) = -\text{Im} \langle \Psi_0(\kappa) | \nabla_\kappa \Psi_0(\kappa) \rangle. \quad (12)$$

Dealing with the disparate manifestations of the Berry phase which occur in molecular and condensed matter phenomena is clearly beyond the scope of the present review; we quote Refs. [18–22] out of many possible references.

Connection and curvature

The loop integral of the Berry connection (i.e. the Berry phase γ) is non trivial in two cases: either the curl of $\mathbf{A}(\kappa)$ is nonzero, or the curl is zero but the curve C is not in a simply connected domain. In the former case, we can invoke Stokes theorem; the formulation is very simple when κ is a 3d parameter. If C is the boundary of a surface Σ (i.e. $C \equiv \partial\Sigma$), and the curl of $\mathbf{A}(\kappa)$ is regular on Σ , then Stokes’ theorem reads

$$\gamma = \oint_{\partial\Sigma} \mathbf{A}(\kappa) \cdot d\kappa = \int_\Sigma \boldsymbol{\Omega}(\kappa) \cdot \mathbf{n} d\sigma, \quad (13)$$

where $\boldsymbol{\Omega}$ is the Berry curvature, defined as

$$\begin{aligned} \boldsymbol{\Omega}(\kappa) &= \nabla_\kappa \times \mathbf{A}(\kappa) = -\text{Im} \langle \nabla_\kappa \Psi_0(\kappa) | \times | \nabla_\kappa \Psi_0(\kappa) \rangle \\ &= i \langle \nabla_\kappa \Psi_0(\kappa) | \times | \nabla_\kappa \Psi_0(\kappa) \rangle, \end{aligned} \quad (14)$$

with the usual meaning of the cross product between three-component bra and ket states. Equation (13) may be spelled out by saying that the curvature is the Berry phase per unit area of Σ .

For $d \neq 3$ the Berry curvature is conveniently written as the $d \times d$ antisymmetric matrix

$$\Omega_{\alpha\beta}(\boldsymbol{\kappa}) = -2 \operatorname{Im} \langle \partial_\alpha \Psi_0(\boldsymbol{\kappa}) | \partial_\beta \Psi_0(\boldsymbol{\kappa}) \rangle; \quad (15)$$

Greek subscripts are Cartesian coordinates throughout, and $\partial_\alpha = \partial/\partial\kappa_\alpha$. The Stokes theorem can still be applied, generalizing Eq. (13) to

$$\gamma = \frac{1}{2} \int_{\Sigma} d\kappa^\alpha \wedge d\kappa^\beta \Omega_{\alpha\beta}(\boldsymbol{\kappa}). \quad (16)$$

The Berry connection is also known as “gauge potential”, and the Berry curvature as “gauge field” [20]. It is worth pointing out that the former is gauge-dependent, while the latter is gauge-invariant and therefore corresponds in general to a measurable quantity, even before any integration. The two quantities play (in $\boldsymbol{\kappa}$ -space) a similar role as the vector potential and the magnetic field in elementary magnetostatics: $\mathbf{A}(\mathbf{r})$ is gauge-dependent, nonmeasurable; $\mathbf{B}(\mathbf{r}) = \nabla_{\mathbf{r}} \times \mathbf{A}(\mathbf{r})$ is gauge-invariant, measurable.

The Berry phase γ , defined as the integral over a closed curve C of the connection, is gauge invariant only modulo 2π . This indeterminacy is resolved by Eqs. (13) and (16) whenever the curve C is the boundary $\partial\Sigma$ of a surface Σ where the curvature is regular. In fact, the curvature is gauge-invariant and has no modulo 2π indeterminacy.

Chern number

The rhs of Eqs. (13) and (16) is the flux of the Berry curvature on the surface Σ ; such flux remains meaningful even on a closed surface (e.g. a sphere or a torus), in which case $\partial\Sigma$ is the empty set. The key result is that such an integral is quantized. Here we limit ourselves to 3d where

$$\int_{\Sigma} \boldsymbol{\Omega}(\boldsymbol{\kappa}) \cdot \mathbf{n} d\sigma = 2\pi C_1; \quad (17)$$

C_1 is an integer, called Chern number of the first class.

The proof is based on a similar algebra as for Dirac’s theory of the magnetic monopole [19, 23], and is also closely related to the Gauss-Bonnet theorem in differential geometry.

The curvature is regular (and divergence-free) on the closed surface Σ ; the lhs of Eq. (17) is the flux of $\boldsymbol{\Omega}(\boldsymbol{\kappa})$ across Σ . The integrand $\boldsymbol{\Omega}(\boldsymbol{\kappa})$ is the curl of the connection $\mathbf{A}(\boldsymbol{\kappa})$; the latter in general *cannot* be defined as a single-valued function globally on Σ , but only on *patches* of it [19, 23]. To fix the ideas, suppose that $\boldsymbol{\Omega}(\boldsymbol{\kappa})$ is singular at $\boldsymbol{\kappa} = 0$, and that Σ is a spherical

surface centered at the origin. We cut this surface at the equator $\kappa_z = 0$ and we consider the flux across the two open surfaces:

$$\int_{\Sigma} \boldsymbol{\Omega}(\boldsymbol{\kappa}) \cdot \mathbf{n} d\sigma = \int_{\Sigma_+} \boldsymbol{\Omega}(\boldsymbol{\kappa}) \cdot \mathbf{n} d\sigma + \int_{\Sigma_-} \boldsymbol{\Omega}(\boldsymbol{\kappa}) \cdot \mathbf{n} d\sigma. \quad (18)$$

We notice that $\partial\Sigma_+ = \partial\Sigma_- = C$, but the surface normals \mathbf{n} have opposite orientations. From Stokes theorem, Eq. (13), we get:

$$\int_{\Sigma_{\pm}} \boldsymbol{\Omega}(\boldsymbol{\kappa}) \cdot \mathbf{n} d\sigma = \pm \oint_C \mathbf{A}_{\pm}(\boldsymbol{\kappa}) \cdot d\boldsymbol{\kappa} \quad (19)$$

$$\int_{\Sigma} \boldsymbol{\Omega}(\boldsymbol{\kappa}) \cdot \mathbf{n} d\sigma = \oint_C \mathbf{A}_+(\boldsymbol{\kappa}) \cdot d\boldsymbol{\kappa} - \oint_C \mathbf{A}_-(\boldsymbol{\kappa}) \cdot d\boldsymbol{\kappa}. \quad (20)$$

The two upper and lower Berry connections $\mathbf{A}_{\pm}(\boldsymbol{\kappa})$ may only differ by a gauge transformation; the rhs of Eq. (20) is the difference of two Berry phases on the same path and is necessarily a multiple of 2π . This concludes the proof of Eq. (17).

We emphasize that the Chern number is a robust topological invariant of the wavefunction, and is at the origin of observable effects. The Chern number made its first appearance in electronic structure in 1982, in the famous TKNN paper about the quantum Hall effect [24]. Nowadays more complex topological invariants are in fashion, and they characterize a completely novel class of insulators, called “topological insulators” [25–31].

Metric

Starting from Eq. (5), the infinitesimal distance is

$$D_{\boldsymbol{\kappa}, \boldsymbol{\kappa} + d\boldsymbol{\kappa}}^2 = \sum_{\alpha, \beta=1}^d g_{\alpha\beta}(\boldsymbol{\kappa}) d\kappa_\alpha d\kappa_\beta, \quad (21)$$

where the metric tensor is easily shown to be

$$\begin{aligned} g_{\alpha\beta}(\boldsymbol{\kappa}) &= \operatorname{Re} \langle \partial_\alpha \Psi_0(\boldsymbol{\kappa}) | \partial_\beta \Psi_0(\boldsymbol{\kappa}) \rangle \\ &- \langle \partial_\alpha \Psi_0(\boldsymbol{\kappa}) | \Psi_0(\boldsymbol{\kappa}) \rangle \langle \Psi_0(\boldsymbol{\kappa}) | \partial_\beta \Psi_0(\boldsymbol{\kappa}) \rangle \\ &= \operatorname{Re} \langle \partial_\alpha \Psi_0(\boldsymbol{\kappa}) | \hat{Q}(\boldsymbol{\kappa}) | \partial_\beta \Psi_0(\boldsymbol{\kappa}) \rangle; \end{aligned} \quad (22)$$

the projector $\hat{Q}(\boldsymbol{\kappa})$ is the same as defined in Eq. (2). This quantum metric tensor was first proposed by Provost and Vallee in 1980 [32].

At this point we may compare Eq. (22) to Eq. (15), noticing that the insertion of $\hat{Q}(\boldsymbol{\kappa})$ is irrelevant in the latter, i.e.

$$\Omega_{\alpha\beta}(\boldsymbol{\kappa}) = -2 \operatorname{Im} \langle \partial_\alpha \Psi_0(\boldsymbol{\kappa}) | \hat{Q}(\boldsymbol{\kappa}) | \partial_\beta \Psi_0(\boldsymbol{\kappa}) \rangle. \quad (23)$$

It is therefore clear that $g_{\alpha\beta}$ and $\Omega_{\alpha\beta}$ are, apart for a trivial -2 factor, the real (symmetric) and the imaginary

(antisymmetric) parts of the same tensor, which we are going to call $\eta_{\alpha\beta}$ in the following:

$$\eta_{\alpha\beta}(\kappa) = \langle \partial_\alpha \Psi_0(\kappa) | \hat{Q}(\kappa) | \partial_\beta \Psi_0(\kappa) \rangle. \quad (24)$$

The curvature, the metric, and hence the full tensor $\eta_{\alpha\beta}$ are all gauge-invariant quantities. A compact equivalent expression is

$$\eta_{\alpha\beta}(\kappa) = \text{Tr} \{ \partial_\alpha \hat{P}(\kappa) \hat{Q}(\kappa) \partial_\beta \hat{P}(\kappa) \}, \quad (25)$$

manifestly gauge-invariant and Hermitian.

Sum over states

The κ -derivatives entering many of the previous equations, e.g. Eq. (24), can be expressed starting from perturbation theory:

$$\begin{aligned} & |\Psi_0(\kappa + \Delta\kappa)\rangle - |\Psi_0(\kappa)\rangle \\ & \simeq \sum'_{n \neq 0} |\Psi_n(\kappa)\rangle \frac{\langle \Psi_n(\kappa) | [H(\kappa + \Delta\kappa) - H(\kappa)] | \Psi_0(\kappa) \rangle}{E_0(\kappa) - E_n(\kappa)}; \end{aligned} \quad (26)$$

$$|\partial_\alpha \Psi_0(\kappa)\rangle = \sum'_{n \neq 0} |\Psi_n(\kappa)\rangle \frac{\langle \Psi_n(\kappa) | \partial_\alpha H(\kappa) | \Psi_0(\kappa) \rangle}{E_0(\kappa) - E_n(\kappa)}. \quad (27)$$

These seemingly obvious and innocent formulae need some caveat. It is clear that inserting Eq. (27) into the Berry connection, Eqs. (11) and (12), we get a vanishing result for any κ . This happens because the simple expression of Eq. (26) corresponds to a very specific gauge choice (called the parallel-transport gauge [21]); multiplying the rhs by a κ -dependent phase factor is legitimate, and must not modify any physical result, while the Berry connection is instead affected. Nonetheless, since our $\eta_{\alpha\beta}(\kappa)$ is a gauge-invariant quantity, we may safely evaluate it in any gauge, including the parallel-transport gauge, implicit in Eq. (27). The result is

$$\begin{aligned} & \eta_{\alpha\beta}(\kappa) \\ & = \sum'_{n \neq 0} \frac{\langle \Psi_0(\kappa) | \partial_\alpha H(\kappa) | \Psi_n(\kappa) \rangle \langle \Psi_n(\kappa) | \partial_\beta H(\kappa) | \Psi_0(\kappa) \rangle}{[E_0(\kappa) - E_n(\kappa)]^2}. \end{aligned} \quad (28)$$

This expression shows explicitly that both the curvature and the metric are ill defined and singular wherever the ground state is degenerate with the first excited state. Indeed, this is the main reason why the domain may happen not to be simply connected.

MANY-ELECTRON WAVEFUNCTION

In the previous Section we have discussed some geometrical aspects of quantum mechanics, remaining

at a rather abstract level, and without addressing either specific physical systems, or measurable quantities. From now on, instead, we address a many-electron system, whose most general Hamiltonian we write, in the Schrödinger representation and in Gaussian units, as

$$\hat{H}(\kappa) = \frac{1}{2m_e} \sum_{i=1}^N |\mathbf{p}_i + \frac{e}{c} \mathbf{A}(\mathbf{r}_i) + \hbar\kappa|^2 + \hat{V}, \quad (29)$$

where m_e and $-e$ are the electron mass and charge, respectively ($e > 0$). Equation (29) is exact in the nonrelativistic, infinite-nuclear-mass limit. In Eq. (29) $\mathbf{A}(\mathbf{r})$ is a vector potential of magnetic origin and the potential \hat{V} includes one-body and two-body (electron-electron) contributions. The 3d parameter κ , having the dimensions of an inverse length, can be regarded as a pure gauge, constant in space, which does not affect the fields (electric and magnetic) and therefore is irrelevant in classical mechanics. Matters are different in quantum mechanics, where the *potentials* (vector and scalar) are the most fundamental quantities, as we know well since the Bohm-Aharonov paper in 1959 [33, 34]. In our case the parameter κ , called “flux” or “twist”, affects indeed the eigenvalues and eigenvectors depending on the chosen boundary conditions.

Boundary conditions

Two very different kinds of boundary conditions occur in electronic structure theory: “open” boundary conditions (OBCs) and periodic boundary conditions (PBCs). The former are appropriate to molecular physics, and require that the many-electron wavefunction of a bound state is square-integrable over the whole coordinate space \mathbb{R}^{3N} . PBCs are instead appropriate for extended systems, either crystalline or disordered, either independent-electron or correlated. For the sake of simplicity here we adopt PBCs over a cubic box of side L , meaning that the eigenstates of Eq. (29) at any given κ are Born-von-Kàrmà̄n periodic in the cubic box over each electron coordinate independently; their Cartesian components $r_{i,\alpha}$ are then equivalent to the angles $2\pi r_{i,\alpha}/L$. The potential \hat{V} enjoys the same periodicity, which implies that the microscopic electric field averages to zero over the sample. The eigenfunctions are normalized in the hypercube of volume L^{3N} .

As previously observed, the choice of boundary conditions (either OBCs or PBCs) corresponds to choosing the Hilbert space, which in turn affects profoundly the geometrical properties discussed in Sec. .

Open boundary conditions

The case of OBCs is by far the simplest. We write for the sake of simplicity the ground state of Eq. (29) at $\kappa = 0$ as $|\Psi_0\rangle \equiv |\Psi_0(0)\rangle$, and we define the many-body position operator as

$$\hat{\mathbf{r}} = \sum_{i=1}^N \mathbf{r}_i. \quad (30)$$

It is straightforward to verify that the state $e^{-i\kappa \cdot \hat{\mathbf{r}}} |\Psi_0\rangle$ fulfills both the Schrödinger equation at the given κ and the boundary conditions with a κ -independent eigenvalue. In jargon we say that the κ -dependence is easily “gauged away” within OBCs; we anticipate that matters are quite different within PBCs.

The state $e^{-i\kappa \cdot \hat{\mathbf{r}}} |\Psi_0\rangle$ coincides therefore with the ground eigenstate $|\Psi_0(\kappa)\rangle$ of the twisted Hamiltonian, Eq. (29). It is legitimate to multiply this eigenvector by any κ -dependent (and position-independent) phase factor; our choice is then

$$|\Psi_0(\kappa)\rangle = e^{-i\kappa \cdot (\hat{\mathbf{r}} - \mathbf{d})} |\Psi_0\rangle, \quad (31)$$

where $\mathbf{d} = \langle \Psi_0 | \hat{\mathbf{r}} | \Psi_0 \rangle$ is the electronic dipole of the molecular system. It follows that the κ -derivative needed in our geometrical quantities (curvature and metric) is

$$|\nabla_\kappa \Psi_0\rangle = -i(\hat{\mathbf{r}} - \mathbf{d}) |\Psi_0\rangle = -i\hat{Q}(0) \hat{\mathbf{r}} |\Psi_0\rangle, \quad (32)$$

Exploiting the idempotency of the projector, the tensor $\eta_{\alpha\beta}$, evaluated at $\kappa = 0$, is

$$\begin{aligned} \eta_{\alpha\beta}(0) &= \langle \Psi_0 | \hat{r}_\alpha \hat{Q}(0) \hat{r}_\beta | \Psi_0 \rangle \\ &= \langle \Psi_0 | \hat{r}_\alpha \hat{r}_\beta | \Psi_0 \rangle - \langle \Psi_0 | \hat{r}_\alpha | \Psi_0 \rangle \langle \Psi_0 | \hat{r}_\beta | \Psi_0 \rangle \end{aligned} \quad (33)$$

This is clearly a real symmetric tensor, hence the Berry curvature vanishes within OBCs, and the metric $g_{\alpha\beta}(0)$ coincides with $\eta_{\alpha\beta}(0)$. It will be shown that within PBCs, instead, the tensor $\eta_{\alpha\beta}(0)$ may acquire a nonvanishing imaginary part.

Clearly, the metric tensor in Eq. (33) is the cumulant second moment of the position operator, or equivalently the ground state fluctuation of the dipole of the molecular system; this quantity is extensive (scales like N in macroscopically homogenous systems). We anticipate that $\eta_{\alpha\beta}(0)/N$ discriminates, in the large N -limit, between insulators and metals.

Periodic boundary conditions

First of all, some preliminary observations about the position operator are in order. The simple multiplicative operator $\hat{\mathbf{r}}$, as defined in Eq. (30), is “forbidden” within PBCs: in fact it maps any periodic wavefunction $|\Psi\rangle$ into

the *nonperiodic* function $\hat{\mathbf{r}}|\Psi\rangle$. In other words, it maps a function which belongs to the Hilbert space to a function outside of it [35]. Incidentally, this is the main reason why the polarization problem has remained unsolved until the early 1990s (for historical presentations, see e.g Refs. [10–12]). In the present context, formulae like Eq. (33) are ill defined and absurd within PBCs.

Suppose that $|\Psi_0\rangle$ is the ground eigenstate of Eq. (29) at $\kappa = 0$. Then the simple state $e^{-i\kappa \cdot \hat{\mathbf{r}}} |\Psi_0\rangle$ fulfills the Schrödinger equation but *not*—for a general κ —PBCs. It is therefore not an eigenstate, except for the discrete κ values

$$\kappa_{m_1 m_2 m_3} = \frac{2\pi}{L} (m_1 \mathbf{e}_1 + m_2 \mathbf{e}_2 + m_3 \mathbf{e}_3), \quad (34)$$

where $m_\alpha \in \mathbb{Z}$, and \mathbf{e}_α are the Cartesian versors. For fractional values of κ —i.e. for values different from those in Eq. (34)—the κ -dependence of the eigenvalues and eigenvectors of Eq. (29) is nontrivial.

If $|\Psi_0(\kappa)\rangle$ is the genuine ground eigenstate of Eq. (29) within PBCs, then the auxiliary function $|\tilde{\Psi}_0(\kappa)\rangle = e^{i\kappa \cdot \hat{\mathbf{r}}} |\Psi_0(\kappa)\rangle$ is a solution of $\hat{H}(0)$, but fulfills quasi-periodic boundary conditions: at any two opposite faces of the cube the wavefunction differs by a κ -dependent phase factor. In other words the problem can be formulated in two equivalent ways: either the Hamiltonian is κ -dependent, as in Eq. (29), and the boundary conditions are κ -independent; or the Hamiltonian is κ -independent but the boundary conditions are “twisted” in a κ -dependent way.

We briefly switch to a 1d formulation in order to explain the role of κ , and even the alternative semantics (“flux” instead of “twist”). Setting the magnetic vector potential to zero, the 1d version of Eq. (29) is

$$\hat{H}(\kappa) = \frac{1}{2m_e} \sum_{i=1}^N (p_i + \hbar\kappa)^2 + \hat{V}, \quad (35)$$

and we require the wavefunction to be periodical with period L over each electron coordinate independently. Alternatively, we may regard this problem as if the electrons were confined to a circular rail of circumference L (see Fig. 4). There is no fields (electric or magnetic) on the rail, but the Hamiltonian is the same as if a constant vector potential of intensity $A = c\hbar\kappa/e$ were present along the rail. This corresponds to a magnetic flux $\Phi = c\hbar\kappa L/e$ piercing the surface encircled by the rail, in a region *not visited* by the electronic system; it has been appropriately called by some authors “inaccessible flux”. This flux affects nontrivially the eigenvalues and eigenvectors, as it is easily verified in the trivial case $\hat{V} \equiv 0$. The inaccessible flux has therefore observable effects; a similar algebra in a different context leads indeed to the Bohm-Aharonov effect [33, 34]. When κ assumes one of the values in Eq. (34), then the flux Φ is an integer multiple of the elementary flux quantum

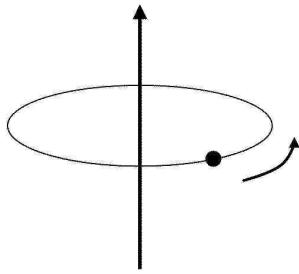


FIG. 4: The electron motion is confined to a circular rail. A constant vector potential along the rail, as in Eq. (35), corresponds to vanishing fields (electric and magnetic), yet the spectrum *depends* on the “inaccessible flux” piercing the surface encircled by the rail.

$\Phi_0 = 2\pi\hbar c/e = hc/e$ ($\Phi_0 = h/e$ in SI units). We stress that only the fractional part of the flux affects the results in a nontrivial way.

It has already been observed that within PBCs the tensor $\eta_{\alpha\beta}(0)$ cannot be simplified—as e.g. in Eq. (33)—and must be therefore addressed in its original form, Eq. (24), by actually evaluating the κ derivatives. As within OBCs, even within PBCs these tensors are extensive. In the thermodynamic limit (i.e $N \rightarrow \infty$, keeping N/L^3 constant), the well defined quantity is $\eta_{\alpha\beta}(0)/N$.

Finally, we observe that for time-reversal symmetric systems $\eta_{\alpha\beta}(0)$ is real symmetric, and coincides therefore with the metric tensor $g_{\alpha\beta}(0)$.

THE LOCALIZATION TENSOR

In the milestone paper, appeared in 1964 under the title “Theory of the insulating state” [7, 8], W. Kohn addressed the insulating character of a solid from a very general viewpoint. Kohn focusses on the electronic *ground state* as a whole. His outstanding message is that “insulating behavior does not appear to depend on the notion of a filled band but rather reflects a certain type of organization of the electrons”. Furthermore, the electrons “so organize themselves as to satisfy a many electron localization condition” [8]. Such localization, however, must be defined in a subtle way. For instance, the (single-particle) Hamiltonian eigenstates of a band insulator are definitely *not* localized.

The modern formulations of the theory of the insulating state are all based on a localization tensor (squared localization length in 1d), first introduced by Resta and Sorella in 1999 [13], as the main tool to address the “organization of the electrons” in the ground state. This work was followed soon afterwards by SWM [14], where additional results are found, most notably relating localization to conductivity. Since then, most authors (including the present one) have adopted the

notation $\langle r_\alpha r_\beta \rangle_c$ for the localization tensor, where “c” stays for cumulant. The formulation within OBCs, using the language and the notations familiar in quantum chemistry, dates since 2006 [36].

The tensor $\langle r_\alpha r_\beta \rangle_c$ has the dimensions of a squared length; it is an intensive quantity that characterizes the ground-state many-body wavefunction as a whole. Its key virtue is that it discriminates between insulators and metals: it is finite in the former case and divergent (in the large-system limit) in the latter. This is the main message of the present work (and of the modern theory of the insulating state): we are going to prove it below, Sect.

Several definitions of the localization tensor have been given in the literature, all of them equivalent. Here we give a novel one, deeply rooted in the geometrical concepts discussed so far: the localization tensor is the intensive quantity

$$\langle r_\alpha r_\beta \rangle_c = \eta_{\alpha\beta}(0)/N, \quad (36)$$

where the thermodynamic limit is understood. A glance at the OBCs expression, Eq. (33), explains the reason for the notation, which we adopt within PBCs as well.

Until 2005 the theory of the insulating state implicitly addressed time-reversal invariant systems only, where the η tensor is real symmetric, and coincides with the metric: in such systems therefore

$$\langle r_\alpha r_\beta \rangle_c = g_{\alpha\beta}(0)/N. \quad (37)$$

It was found in 2005 [37] that—in absence of time-reversal symmetry and within PBCs—the tensor $\langle r_\alpha r_\beta \rangle_c$ is naturally endowed with an antisymmetric imaginary part, whose physical meaning is also outstanding. This is discussed in Sects. and .

Discretized formulae

While practical implementations within OBCs invariably use Eq. (33), those within PBCs require a discretization in κ -space in order to evaluate the derivatives appearing in Eq. (24). A key feature is that any discretization must be *numerically* gauge invariant. Here we are going to present a novel discretized formula, based on Eq. (25) at $\kappa = 0$.

By definition we have

$$\partial_\alpha \hat{P}(0) = \lim_{k \rightarrow 0} \frac{1}{k} [\hat{P}(k\mathbf{e}_\alpha) - \hat{P}(0)], \quad (38)$$

hence for small k

$$\begin{aligned} \eta_{\alpha\beta}(0) &\simeq \frac{1}{k^2} \text{Tr} \{ [\hat{P}(k\mathbf{e}_\alpha) - \hat{P}(0)] \hat{Q}(0) [\hat{P}(k\mathbf{e}_\beta) - \hat{P}(0)] \} \\ &= \frac{1}{k^2} \text{Tr} \{ \hat{P}(k\mathbf{e}_\alpha) \hat{Q}(0) \hat{P}(k\mathbf{e}_\beta) \} \end{aligned} \quad (39)$$

$$\begin{aligned}
&= \frac{1}{k^2} \langle \Psi_0(k\mathbf{e}_\alpha) | \Psi_0(k\mathbf{e}_\beta) \rangle \langle \Psi_0(k\mathbf{e}_\beta) | \Psi_0(k\mathbf{e}_\alpha) \rangle \\
&\quad - \frac{1}{k^2} \langle \Psi_0(k\mathbf{e}_\alpha) | \Psi_0 \rangle \langle \Psi_0 | \Psi_0(k\mathbf{e}_\beta) \rangle \langle \Psi_0(k\mathbf{e}_\beta) | \Psi_0(k\mathbf{e}_\alpha) \rangle.
\end{aligned} \tag{40}$$

The gauge invariance is pescuous: all the arbitrary phase factors cancel in pairs.

Equation (40) is still a three-point formula; it is transformed to a single-point one using a key result from Ref. [13]. Whenever k is equal to $\frac{2\pi}{L}$ times an integer, the wavefunction $\mathbf{e}^{-ik\hat{r}_\alpha} |\Psi_0(0)\rangle$ obeys PBCs—see also Eq. (34)—and is an eigenstate of $\hat{H}(k\mathbf{e}_\alpha)$ with eigenvalue E_0 . For insulating systems the ground state is nondegenerate even in the thermodynamic limit, hence

$$|\Psi_0(\frac{2\pi}{L}\mathbf{e}_\alpha)\rangle = \mathbf{e}^{-i2\pi\hat{r}_\alpha/L} |\Psi_0\rangle, \tag{41}$$

apart possibly by a phase factor, irrelevant here. Replacing k with $\frac{2\pi}{L}$ for large L in Eq. (40), and using Eq. (41) we get the sought for formula, where a single diagonalization of Eq. (29) at $\kappa = 0$ is enough to obtain both the real and imaginary part of the localization tensor $\langle r_\alpha r_\beta \rangle_c$. Notice that the first line of Eq. (40) is real.

The formulae previous reported in the literature are inspired by the Berry-phase formulae—see Eq. (7)—and based on logarithms. To retrieve them, it is enough to use $\log(1 + \varepsilon) \simeq \epsilon$ in Eq. (40):

$$\eta_{\alpha\beta}(0) \simeq \frac{1}{k^2} \log \frac{\langle \Psi_0(k\mathbf{e}_\alpha) | \Psi_0(k\mathbf{e}_\beta) \rangle}{\langle \Psi_0(k\mathbf{e}_\alpha) | \Psi_0 \rangle \langle \Psi_0 | \Psi_0(k\mathbf{e}_\beta) \rangle}, \tag{42}$$

which provides $\langle r_\alpha r_\beta \rangle_c$ identical in form to Eq. (86) in Ref. [15]; for $\alpha = \beta$ the formula is also identical to the archetypical one, proposed by Resta and Sorella in 1999 [13]. For the real part of $\langle r_\alpha r_\beta \rangle_c$ the logarithm formula, Eq. (42), has some advantages over Eq. (40). For the imaginary part, instead, the second line of Eq. (40) provides a much better result, since it is *not* affected by any modulo 2π indeterminacy. Getting rid of such indeterminacy, in fact, is essential when dealing with Chern numbers (Sects. and).

Sum over states again

We express $\langle r_\alpha r_\beta \rangle_c$ using the sum-over-states formula, Eq. (28) :

$$\langle r_\alpha r_\beta \rangle_c = \frac{1}{N} \sum'_{n \neq 0} \frac{\langle \Psi_0 | \partial_\alpha \hat{H}(0) | \Psi_n \rangle \langle \Psi_n | \partial_\beta \hat{H}(0) | \Psi_0 \rangle}{(E_0 - E_n)^2}. \tag{43}$$

The κ -derivative of the Hamiltonian of Eq. (29) is

$$\nabla_\kappa \hat{H}(0) = \frac{\hbar}{m_e} \sum_{i=1}^N \left[\mathbf{p}_i + \frac{e}{c} \mathbf{A}(\mathbf{r}_i) \right], \tag{44}$$

where the rhs is nothing else than \hbar times the velocity operator $\hat{\mathbf{v}}$; Eq. (43) becomes then

$$\begin{aligned}
\langle r_\alpha r_\beta \rangle_c &= \frac{1}{\hbar^2 N} \sum'_{n \neq 0} \frac{\langle \Psi_0 | \hat{v}_\alpha | \Psi_n \rangle \langle \Psi_n | \hat{v}_\beta | \Psi_0 \rangle}{(E_0 - E_n)^2} \\
&= \frac{1}{N} \sum'_{n \neq 0} \frac{\langle \Psi_0 | \hat{v}_\alpha | \Psi_n \rangle \langle \Psi_n | \hat{v}_\beta | \Psi_0 \rangle}{\omega_{0n}^2},
\end{aligned} \tag{45}$$

where $\omega_{0n} = (E_n - E_0)/\hbar$.

The velocity operator is also commonly expressed as $\hat{\mathbf{v}} = i[\hat{H}(0), \hat{\mathbf{r}}]/\hbar$, but it is worth emphasizing that while the position $\hat{\mathbf{r}}$ is well defined within OBCs and ill defined within PBCs, the velocity $\hat{\mathbf{v}}$ is well defined in both cases.

The basic sum-over states formula, Eq. (45), applies therefore to both OBCs and PBCs. In general, it is not very useful on practical grounds, since it would require the evaluation of slowly convergent sums. Nonetheless, Eq. (45) is instead essential to gather understanding into the physical meaning of $\eta_{\alpha\beta}$, as will be shown below.

CONDUCTIVITY

Linear response

So far, we have only discussed ground-state properties of our N -electron system. Suppose now that it is subject to a small time-dependent perturbation contributing to the Hamiltonian the term:

$$\delta \hat{H}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega f(\omega) \frac{1}{2} (\hat{A} e^{-i\omega t} + \hat{A}^\dagger e^{i\omega t}), \tag{46}$$

where \hat{A} determines the “shape” of the perturbation and f its amplitude. In order to get an Hermitian $\delta \hat{H}$, we assume $f(\omega) = f(-\omega)$. We wish to measure the response to such perturbation by means of the expectation value of some observable \hat{B} , i.e.:

$$\delta \langle \hat{B} \rangle = \langle \tilde{\Psi} | \hat{B} | \tilde{\Psi} \rangle - \langle \Psi_0 | \hat{B} | \Psi_0 \rangle, \tag{47}$$

where $\tilde{\Psi} = \Psi_0 + \delta\Psi(t)$ is the perturbed time-evolved ground state. If we limit ourselves to study terms which are linear in the response, it is enough to consider the single oscillatory perturbation:

$$\hat{H}'(\omega) = \frac{1}{2} (\hat{A} e^{-i\omega t} + \hat{A}^\dagger e^{i\omega t}), \tag{48}$$

whose response can be written, using the compact notations due to Zubarev [38–40], as:

$$\delta \langle \hat{B} \rangle = \frac{1}{2} (\langle \langle \hat{B} | \hat{A} \rangle \rangle_\omega e^{-i\omega t} + \langle \langle \hat{B} | \hat{A} \rangle \rangle_{-\omega} e^{i\omega t}). \tag{49}$$

The quantity $\langle \langle \hat{B} | \hat{A} \rangle \rangle_\omega$ is by definition the linear response induced by the perturbation \hat{A} at frequency ω on

the expectation value $\langle \hat{B} \rangle$. Straightforward first-order perturbation theory provides its explicit expression as:

$$\begin{aligned} \langle\langle \hat{B}|\hat{A}\rangle\rangle_\omega &= \frac{1}{\hbar} \lim_{\eta \rightarrow 0+} \sum'_{n \neq 0} \left(\frac{\langle \Psi_0 | \hat{B} | \Psi_n \rangle \langle \Psi_n | \hat{A} | \Psi_0 \rangle}{\omega - \omega_{0n} + i\eta} \right. \\ &\quad \left. - \frac{\langle \Psi_0 | \hat{A} | \Psi_n \rangle \langle \Psi_n | \hat{B} | \Psi_0 \rangle}{\omega + \omega_{0n} + i\eta} \right), \end{aligned} \quad (50)$$

where ω_{0n} are the excitation frequencies of the unperturbed system, and the positive infinitesimal η ensures causality. Expressions of the kind of Eq. (50) go under the name of Kubo formulae.

Kubo formula for conductivity

The conductivity tensor $\sigma_{\alpha\beta}(\omega)$ measures the current linearly induced by an electric field: $j_\alpha = \sigma_{\alpha\beta} \mathcal{E}_\beta$. We therefore identify \hat{A} with the potential of an electric field along β , i.e. $\hat{A} = e\mathcal{E}\hat{r}_\beta$, and \hat{B} with the current operator $-e\hat{v}_\alpha/L^3$. An important detail must be stressed at this point. The macroscopic field inside the sample includes by definition screening effects due to the electronic system, while the perturbation $\delta\hat{H}$ entering Eq. (29)—via the \hat{A} operator—is the “bare”, or unscreened one. This point will be discussed below (Sect.); for the time being we simply identify screened and unscreened fields.

The Kubo formula for conductivity is therefore

$$\sigma_{\alpha\beta}(\omega) = -\frac{e^2}{L^3} \langle\langle \hat{v}_\alpha | \hat{r}_\beta \rangle\rangle_\omega; \quad (51)$$

this is correct within OBCs, but meaningless within PBCs, owing to the explicit presence of the position operator. This, however, makes no harm, since only its off-diagonal matrix elements are required: see Eq. (50). As usual, we may exploit the identity $\langle \Psi_0 | \hat{\mathbf{r}} | \Psi_n \rangle = i \langle \Psi_0 | \hat{\mathbf{v}} | \Psi_n \rangle / \omega_{0n}$. The Kubo formula becomes then

$$\begin{aligned} \sigma_{\alpha\beta}(\omega) &= \frac{ie^2}{\hbar L^3} \lim_{\eta \rightarrow 0+} \sum'_{n \neq 0} \frac{1}{\omega_{0n}} \left(\frac{\langle \Psi_0 | \hat{v}_\alpha | \Psi_n \rangle \langle \Psi_n | \hat{v}_\beta | \Psi_0 \rangle}{\omega - \omega_{0n} + i\eta} \right. \\ &\quad \left. + \frac{\langle \Psi_0 | \hat{v}_\beta | \Psi_n \rangle \langle \Psi_n | \hat{v}_\alpha | \Psi_0 \rangle}{\omega + \omega_{0n} + i\eta} \right). \end{aligned} \quad (52)$$

We introduce a compact notation for the real and imaginary parts of the numerators in Eq. (52), i.e.

$$\mathcal{R}_{n,\alpha\beta} = \text{Re} \langle \Psi_0 | \hat{v}_\alpha | \Psi_n \rangle \langle \Psi_n | \hat{v}_\beta | \Psi_0 \rangle, \quad (53)$$

$$\mathcal{I}_{n,\alpha\beta} = \text{Im} \langle \Psi_0 | \hat{v}_\alpha | \Psi_n \rangle \langle \Psi_n | \hat{v}_\beta | \Psi_0 \rangle, \quad (54)$$

which are symmetric and antisymmetric, respectively. Using then

$$\lim_{\eta \rightarrow 0+} \frac{1}{x + i\eta} = \mathcal{P} \frac{1}{x} - i\pi\delta(x), \quad (55)$$

and omitting the principal part, we separate for $\omega > 0$ the symmetric and antisymmetric parts in the conductivity tensor as

$$\begin{aligned} \text{Re } \sigma_{\alpha\beta}^{(+)}(\omega) &= \frac{\pi e^2}{\hbar L^3} \sum'_{n \neq 0} \frac{\mathcal{R}_{n,\alpha\beta}}{\omega_{0n}} \delta(\omega - \omega_{0n}) \\ \text{Re } \sigma_{\alpha\beta}^{(-)}(\omega) &= \frac{2e^2}{\hbar L^3} \sum'_{n \neq 0} \frac{\mathcal{I}_{n,\alpha\beta}}{\omega_{0n}^2 - \omega^2}. \end{aligned} \quad (56)$$

Sum rules

At this point, we are ready to compare with the sum-over-states formulae for the tensor $\langle r_\alpha r_\beta \rangle_c$. In the present notations, we rewrite Eq. (45) as

$$\begin{aligned} \text{Re } \langle r_\alpha r_\beta \rangle_c &= \frac{1}{N} \sum'_{n \neq 0} \frac{\mathcal{R}_{n,\alpha\beta}}{\omega_{0n}^2} \\ \text{Im } \langle r_\alpha r_\beta \rangle_c &= \frac{1}{N} \sum'_{n \neq 0} \frac{\mathcal{I}_{n,\alpha\beta}}{\omega_{0n}^2}. \end{aligned} \quad (57)$$

A glance at Eq. (56) shows that

$$\text{Re } \langle r_\alpha r_\beta \rangle_c = \frac{\hbar L^3}{\pi e^2 N} \int_0^\infty \frac{d\omega}{\omega} \text{Re } \sigma_{\alpha\beta}^{(+)}(\omega) \quad (58)$$

$$\text{Im } \langle r_\alpha r_\beta \rangle_c = \frac{\hbar L^3}{2e^2 N} \text{Re } \sigma_{\alpha\beta}^{(-)}(0). \quad (59)$$

Eq. (58) has been arrived at by SWM in 2000 [14], and Eq. (59) by Resta in 2005 [37]. First of all, these identities show that $\langle r_\alpha r_\beta \rangle_c$, defined here as a basic geometric feature, is indeed a measurable quantity (whenever it does not diverge).

As emphasized throughout this work $\langle r_\alpha r_\beta \rangle_c$ is a ground-state property, while the rhs of Eqs. (58) and (59) are properties of the system *excitations*, owing to the Kubo formula. Indeed, both Eqs. (58) and (59) look like the zero-temperature limit of a fluctuation-dissipation theorem, several forms of which are known in statistical physics [41, 42]: in the lhs we have a ground-state fluctuation—see in particular Eq. (33)—while the ingredient of the rhs is conductivity (dissipation).

Screened vs. unscreened field

The Kubo formula for conductivity has been obtained identifying the \hat{A} operator with $\mathcal{E}\hat{r}_\beta$, where \mathcal{E} is the macroscopic field inside the sample. In general this is not quite correct, since instead $\hat{A} = e\mathcal{E}_0\hat{r}_\beta$, where \mathcal{E}_0 is the “bare” field, i.e. the field that would be present inside the sample *in absence of screening*. The latter originates from the two-body (electron-electron) terms in the potential \hat{V} entering Schrödinger equation. The relationship between \mathcal{E} and \mathcal{E}_0 is not a bulk property,

and depends on the shape of the sample. Alternatively, it depends on the boundary conditions assumed for integrating Poisson equation: we refer to Ref. [12] for a thorough discussion. Whenever $\mathcal{E} \neq \mathcal{E}_0$, the sum rule in Eq. (58) must be modified.

The ground-state fluctuations, as e.g. Eq. (33), are fluctuations of the macroscopic polarization, which in a finite sample induce a surface charge at the boundary. This in turn generates a depolarizing field, which counteracts polarization. Therefore the localization tensor $\langle r_\alpha r_\beta \rangle_c$ depends on the sample shape, or equivalently on the boundary conditions assumed when taking the thermodynamic limit. The choice of PBCs in Eq. (29), however, implies $\mathcal{E} = \mathcal{E}_0$ [43]; hence Eqs. (58) and (59) are correct as they stand within PBCs. This no longer holds within OBCs: in this case Eq. (58) needs to be modified, while $\text{Im } \langle r_\alpha r_\beta \rangle_c = 0$.

Ideally the equality $\mathcal{E} = \mathcal{E}_0$ corresponds to choosing a sample in the form of a slab, and to addressing the component of the fluctuation tensor $\langle r_\alpha r_\beta \rangle_c$ parallel to the slab [12]; the thermodynamic limit amounts then to the infinite slab thickness. Owing to the long range of Coulomb interaction the order of the limits (first a slab, then its infinite thickness) is crucial. For instance, if the limit is taken instead by considering spherical clusters of increasing radius, the SWM fluctuation-dissipation sum rule, Eq. (58), assumes a different form: this is discussed in Refs. [36, 43]. The explicit form of the generalized sum rule is given therein.

Last but not least, the effect leading to $\mathcal{E} \neq \mathcal{E}_0$ within OBCs is a pure correlation effect. It originates from explicitly correlated wavefunctions, and does not occur within mean-field theories (Hartree-Fock and Kohn-Sham) [36, 43]. Within such theories, therefore, the sum rules hold in the simple form of Eqs. (58) and (59); the conductivity therein is the independent-particle conductivity (“uncoupled” response in quantum-chemistry jargon).

LOCALIZATION IN THE INSULATING STATE

The basic tenet of the modern theory of the insulating state is that the localization tensor $\langle r_\alpha r_\beta \rangle_c$ is the ground-state property which sharply discriminates—in the spirit of Kohn’s seminal work [7, 8]—between insulators and metals. The real part of $\langle r_\alpha r_\beta \rangle_c$ remains finite in the thermodynamic limit in any insulator, while it diverges in any metal.

The theory is very general, and has found applications to various different kinds of insulators: band insulators [44–47]; correlated (i.e. Mott) insulators, either by means of Hubbard-like model Hamiltonians [13, 48] or realistic ones [49]; “quantum Hall insulators” [37], and Anderson insulators [50]. As for Chern insulators and topological insulators, no explicit application of the

present theory exists; nonetheless the work of Refs. [51, 52] implicitly shows that even in these cases the ground-state wavefunction is indeed localized in the sense of the present review. Most of these applications are reviewed in Sect. .

The ultimate proof of the key property of $\text{Re } \langle r_\alpha r_\beta \rangle_c$ is based on the SWM sum rule, Eq. (58). Since the tensor is real symmetric, it is enough to consider the diagonal elements (over its principal axes)

$$\text{Re } \langle r_\alpha r_\alpha \rangle_c = \frac{\hbar L^3}{\pi e^2 N} \int_0^\infty \frac{d\omega}{\omega} \text{Re } \sigma_{\alpha\alpha}(\omega). \quad (60)$$

The f -sum rule yields

$$\int_0^\infty d\omega \text{Re } \sigma_{\alpha\alpha}(\omega) = \frac{\omega_p^2}{8} = \frac{\pi e^2 N}{2m_e L^3}, \quad (61)$$

where ω_p is the plasma frequency. Therefore the integral in Eq. (60) always converges at ∞ ; its convergence/divergence is dominated by the small- ω behavior of $\text{Re } \sigma_{\alpha\alpha}(\omega)$.

Suppose first that the spectrum is gapped, i.e. the spacing between the ground state and the first excited state stays finite in the thermodynamic limit. If the gap is E_g the conductivity vanishes for $\omega < E_g/\hbar$, and Eq. (61) yields

$$\begin{aligned} \text{Re } \langle r_\alpha r_\alpha \rangle_c &= \frac{\hbar L^3}{\pi e^2 N} \int_{E_g/\hbar}^\infty \frac{d\omega}{\omega} \text{Re } \sigma_{\alpha\alpha}(\omega) \\ &< \frac{\hbar^2 L^3}{\pi e^2 N E_g} \int_0^\infty d\omega \text{Re } \sigma_{\alpha\alpha}(\omega) \\ &= \frac{\hbar^2}{2m_e E_g}. \end{aligned} \quad (62)$$

This inequality is due to SWM and clearly proves that $\text{Re } \langle r_\alpha r_\beta \rangle_c$ is finite in any gapped insulator, as e.g. band insulators (considered in more detail in Sect.).

The main message of Kohn’s 1964 paper, however, is that “insulating characteristics are a strict consequence of electronic localization (in an appropriate sense) and do not require an energy gap”. For any gapless material, the small- ω behavior of $\text{Re } \sigma_{\alpha\alpha}(\omega)$ is the result of a competition between numerators and denominators in the Kubo formula, Eq. (52). Since we aim at a continuous function of ω , the singularities in Eq. (56) must be smoothed: this can be done by keeping the “dissipation” η finite while performing the thermodynamic limit first [53]. For a band metal the localization tensor diverges (see below). According to SWM, a gapless material is insulating whenever $\text{Re } \sigma_{\alpha\alpha}(\omega) \rightarrow 0$ like a positive power of ω , and metallic otherwise. The only example of gapless insulator considered so far is a model Anderson insulator in 1d [50]. Simulations prove indeed that $\langle x^2 \rangle_c$ is finite therein (Sect.).

Independent electrons

For noninteracting electrons the potential \hat{V} in Eq. (29) is the sum of identical one-body terms: $\hat{V} = \sum_{i=1}^N V(\mathbf{r}_i)$. The many-electron Hamiltonian is separable and the exact ground state $|\Psi_0(\boldsymbol{\kappa})\rangle$ is a Slater determinant of one-particle orbitals (doubly occupied in the singlet case). At a mean-field level, the one-body potential $V(\mathbf{r})$ includes electron-electron interaction in a selfconsistent way. In the Hartree-Fock (HF) framework the Slater determinant is regarded as an approximate many-electron wavefunction. Instead, in the density-functional framework the orbitals—called Kohn-Sham (KS) orbitals—are auxiliary quantities, individually devoid of physical meaning. In particular, their Slater determinant *does not* coincide with the many-electron wavefunction $|\Psi_0(\boldsymbol{\kappa})\rangle$, as a matter of principle. Therefore the exact localization tensor does not coincide, at least in principle, with the one obtained from the Slater determinant of KS orbitals.

So much for the matters of principle. On practical grounds such difference is routinely disregarded (e.g. when dealing with polarization, magnetization [10–12], and more), given that it is not at all clear what is the relative importance of this “intrinsic” error, compared with the errors due to the choice of the functional itself. We therefore address here either the HF or the KS wavefunction $|\Psi_0(\boldsymbol{\kappa})\rangle$, having the form of a Slater determinant.

Whenever the wavefunction is a Slater determinant, all ground-state properties can be explicitly cast in terms of the one-body density matrix

$$\rho(\mathbf{r}, \mathbf{r}') = 2P(\mathbf{r}, \mathbf{r}') = 2 \sum_{j=1}^{N/2} \varphi_j(\mathbf{r}) \varphi_j^*(\mathbf{r}'), \quad (63)$$

where a singlet ground state is assumed, and $\varphi_j(\mathbf{r})$ are the occupied one-particle orbitals (either HF or KS); $P(\mathbf{r}, \mathbf{r}')$ is the projector over the occupied manifold.

The expression for the localization tensor is easily found within OBCs starting from Eq. (33) [36]:

$$\langle r_\alpha r_\beta \rangle_c = \frac{1}{N} \int d\mathbf{r} d\mathbf{r}' (\mathbf{r} - \mathbf{r}')_\alpha (\mathbf{r} - \mathbf{r}')_\beta |P(\mathbf{r}, \mathbf{r}')|^2. \quad (64)$$

If we define the complementary projector

$$Q(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') - P(\mathbf{r}, \mathbf{r}'), \quad (65)$$

an equivalent expression is [44]

$$\langle r_\alpha r_\beta \rangle_c = \frac{2}{N} \text{Tr} \{ r_\alpha P r_\beta Q \}, \quad (66)$$

where “Tr” is the trace over the single-particle Hilbert space (not on Cartesian indices).

If we consider a cluster, cut out of a crystalline solid, Eq. (64) becomes in the large- N limit

$$\langle r_\alpha r_\beta \rangle_c = \frac{1}{N_c} \int_{\text{cell}} d\mathbf{r} \int_{\text{all space}} d\mathbf{r}' (\mathbf{r} - \mathbf{r}')_\alpha (\mathbf{r} - \mathbf{r}')_\beta |P(\mathbf{r}, \mathbf{r}')|^2, \quad (67)$$

where N_c is the number of electrons per crystal cell. According to the discussion in Sec. , we need not to worry about shape issues in taking the limit; we also notice that the density matrix, Eq. (63), is independent of the boundary conditions (either OBCs or PBCs) in the large- N limit.

Band insulators and band metals

As observed, Eq. (67) holds for a crystalline solids. Therefore the inner integral on the rhs must converge for a band insulator, and must diverge for a band metal. This is confirmed by the well known fact that the asymptotic behavior of P is qualitatively different in insulators and in metals. In the former materials, in fact, $P(\mathbf{r}, \mathbf{r}')$ decays exponentially [54–57] for large values of $\mathbf{r} - \mathbf{r}'$: therefore the integral converges and the localization tensor is finite. In conducting materials, instead, $P(\mathbf{r}, \mathbf{r}')$ decays only polynomially, and the inner integral diverges. This divergence can be explicitly verified for the simplest conductor of all, namely, the noninteracting electron gas, whose density matrix is exactly known in analytic form [44, 58]. Therefore the localization tensor, when expressed in the form of Eq. (67), measures in a perspicuous way the “nearsightedness”[59] of the electron distribution. Such measure is *qualitatively* different in insulators and in metals.

The one-particle orbitals (either HF or KS) in a crystalline solid have the Bloch form. We therefore may wish to replace the orbitals in the expression for $P(\mathbf{r}, \mathbf{r}')$, Eq. (63),

$$\varphi_i(\mathbf{r}) \rightarrow \psi_{n\mathbf{q}}(\mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}} u_{n\mathbf{q}}(\mathbf{r}), \quad (68)$$

where n is the band index and \mathbf{q} is the Bloch vector. We stress that PBCs are at the very root of Bloch theorem. If the orbitals are normalized to one over the crystal cell of volume V_c , the ground-state projector in insulating crystals is

$$\begin{aligned} P(\mathbf{r}, \mathbf{r}') &= \frac{V_c}{(2\pi)^3} \sum_{n=1}^{N_c/2} \int_{\text{BZ}} d\mathbf{q} \psi_{n\mathbf{q}}(\mathbf{r}) \psi_{n\mathbf{q}}^*(\mathbf{r}') \\ &= \frac{V_c}{(2\pi)^3} \sum_{n=1}^{N_c/2} \int_{\text{BZ}} d\mathbf{q} e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} u_{n\mathbf{q}}(\mathbf{r}) u_{n\mathbf{q}}^*(\mathbf{r}'), \end{aligned} \quad (69)$$

where $N_c/2$ is the number of occupied bands, and the integral is taken over the Brillouin zone (equivalently, over the reciprocal cell).

Using Eq. (69), the localization tensor in a band insulator becomes (for double occupancy)

$$\begin{aligned} \langle r_\alpha r_\beta \rangle_c = & \frac{2V_c}{(2\pi)^3 N_c} \int_{BZ} d\mathbf{q} \left(\sum_n \langle \frac{\partial}{\partial q_\alpha} u_{n\mathbf{q}} | \frac{\partial}{\partial q_\beta} u_{n\mathbf{q}} \rangle \right. \\ & \left. - \sum_{n,n'} \langle u_{n\mathbf{q}} | \frac{\partial}{\partial q_\alpha} u_{n'\mathbf{q}} \rangle \langle \frac{\partial}{\partial q_\beta} u_{n'\mathbf{q}} | u_{n\mathbf{q}} \rangle \right). \end{aligned} \quad (70)$$

The proof is given in Refs. [15, 44], and will not be repeated here.

Wannier functions

Expressions similar to Eq. (70) also enter the Marzari-Vanderbilt theory of maximally localized Wannier functions [60]; the relationship is

$$\sum_{\alpha=1}^d \langle r_\alpha r_\alpha \rangle_c = \frac{2}{N_c} \Omega_I. \quad (71)$$

Here Ω_I indicates the gauge-invariant part of the quadratic spread of the Wannier functions, as in Ref. [60] and in the subsequent literature.

The trace in Eq. (71) is a lower bound (and *not* a minimum in dimension $d > 1$) for the spherical second (cumulant) moment—a.k.a. quadratic spread—of the Wannier functions, averaged over the sample.

We stress that Eqs. (69) and (70) make sense only insofar the Fermi level falls in a gap, in which case $\langle r_\alpha r_\beta \rangle_c$ is always finite. If we vary the Hamiltonian continuously, allowing the gap to close, then $\langle r_\alpha r_\beta \rangle_c$ diverges; the quadratic spread of the Wannier functions diverges as well [36].

Chern insulators

Within OBCs the localization tensor is always real: this is perspicuous in Eqs. (64) and (67). Instead the PBC expression of Eq. (70) is naturally endowed with an imaginary part:

$$\begin{aligned} \text{Im } \langle r_\alpha r_\beta \rangle_c = & \frac{2V_c}{(2\pi)^3 N_c} \times \\ & \times \int_{BZ} d\mathbf{q} \sum_n \text{Im} \langle \frac{\partial}{\partial q_\alpha} u_{n\mathbf{q}} | \frac{\partial}{\partial q_\beta} u_{n\mathbf{q}} \rangle. \end{aligned} \quad (72)$$

A necessary condition for this to be nonzero is the absence of time-reversal symmetry.

The second line of Eq. (72) has (in 3d) the dimensions of an inverse length, and is quantized in units of π times a reciprocal vector. We call this reciprocal vector “Chern invariant”; in fact it is a Chern number (Sect.) in \mathbf{q} -space and in appropriate units. In 2d the BZ integral is

dimensionless and simply related to the Chern number, defined as:

$$C_1 = -\frac{1}{\pi} \int_{BZ} d\mathbf{q} \sum_n \text{Im} \langle \frac{\partial}{\partial q_1} u_{n\mathbf{q}} | \frac{\partial}{\partial q_2} u_{n\mathbf{q}} \rangle \quad (73)$$

(the formula here is given for single band occupancy).

Band insulators where the Chern invariant is nonzero are called “Chern insulators” (normal insulators otherwise). The possible existence of Chern insulators has been pointed out by Haldane in 1988 [61], by means of a remarkable model Hamiltonian in 2d, much studied afterwards [51, 62, 63], and illustrated below (Sect.). We emphasize that the nonvanishing of the Chern number prevents the existence of exponentially localized Wannier functions [51, 64], at variance with normal insulators where the “maximally localized” Wannier functions [60] are exponentially localized [65].

LOCALIZATION IN DIFFERENT KINDS OF INSULATORS

Small molecules

The modern theory of the insulating state clearly addresses extended systems, i.e. the $N \rightarrow \infty$ limit; indeed it makes little sense to ask whether a small molecule is insulating or conducting. Nonetheless the concepts of localized/delocalized electronic states is of the utmost importance in quantum chemistry as well, notably in relationship to aromaticity.

The tensor $\langle r_\alpha r_\beta \rangle_c$ within OBCs is always real symmetric. If the ground-state wavefunction is a Slater determinant, then the trace of the tensor at finite N has the meaning of a lower bound for the quadratic spread of the Boys localized orbitals, averaged over all the occupied orbitals [36].

The small- N version of the main concepts of the present review (in their OBCs flavour [36]) has been adopted in quantum chemistry by Àngyà [66, 67]. Besides providing HF calculations of $\langle r_\alpha r_\beta \rangle_c$ for a sample of small molecules, Àngyà even provides *experimental values* drawn from compilations of the dipole oscillation-strength distributions: basically, from Eq. (58).

Band insulators

The theory warrants that the localization tensor is finite in any insulator. However, quantitative calculations for both model tight-binding Hamiltonians and realistic solids within density-functional theory have been used to illustrate the theory and to identify trends. For instance one expects much smaller diagonal elements $\langle r_\alpha r_\alpha \rangle_c$ for strong (i.e. large-gap) insulators than for

weak (small-gap) ones. This is also suggested by the SWM inequality, Eq. (62).

Let us start with a simple tight-binding (a.k.a. Hückel) Hamiltonian in 1d:

$$\hat{H} = \sum_{j\sigma} [(-1)^j \Delta c_{j\sigma}^\dagger c_{j\sigma} - t(c_{j\sigma}^\dagger c_{j+1\sigma} + \text{H.c.})] \quad (74)$$

where $t > 0$ is the first neighbor hopping ($\beta = -t$ in most chemistry literature) and H.c. stays for Hermitian conjugate. This toy model schematizes a binary ionic crystal; the band structure is

$$\epsilon(q) = \pm \sqrt{\Delta^2 + 4t^2 \cos^2 qa/2}, \quad (75)$$

where a is the lattice constant and q is the Bloch vector. The gap is equal to 2Δ ; at half filling the system is always insulating except for $\Delta = 0$. The squared localization length (within OBCs) is the tight-binding version of Eq. (64), i.e.

$$\langle x^2 \rangle_c = \frac{a^2}{4N} \sum_{j,j'=1}^N P_{jj'}^2 (j-j')^2. \quad (76)$$

This is a monotonically decreasing function of t/Δ ; it is easily verified that it vanishes in the extreme ionic case ($t = 0$). In the metallic case ($\Delta = 0$) the ground-state projector has a simple analytical form:

$$\begin{aligned} P_{jj} &= \frac{1}{2}; \quad P_{jj'} = 0 \text{ for even } |j' - j| = 2s, \\ P_{jj'} &= \frac{(-1)^s}{\pi(2s+1)} \text{ for odd } |j' - j| = 2s+1, \end{aligned} \quad (77)$$

which clearly implies divergence of Eq. (76). At any finite N within OBCs Eq. (76) leads to a finite $\langle x^2 \rangle_c$ value; however Eq. (77) suggests that in the metallic case $\langle x^2 \rangle_c$ diverges linearly with N . This has been verified by actual simulations, even when $\Delta \neq 0$ but the Fermi level is not in the gap [50].

Other simulations [47] have addressed dimerized chains, i.e. $\Delta = 0$ but alternant hoppings in Eq. (74). While nothing relevant occurs within PBCs, partly filled end states within OBCs at some fillings are at the root of some noticeable features.

The first ab-initio study (in 2001) addressed several elemental and binary cubic semiconductors at the KS level [44]. The tensor is real and isotropic. The computed $\langle x^2 \rangle_c$ (Fig. 5) is smaller than 3 bohr² in all the materials studied: the ground many-body wavefunction is therefore very localized in this class of materials. The SWM inequality was also checked, and found to be well verified using both the theoretical KS gap and the experimental one (the latter is typically larger).

Other studies have addressed the ferroelectric perovskites in their different (cubic and noncubic) structures [45], and some model Hamiltonians in 1d and 2d [46].

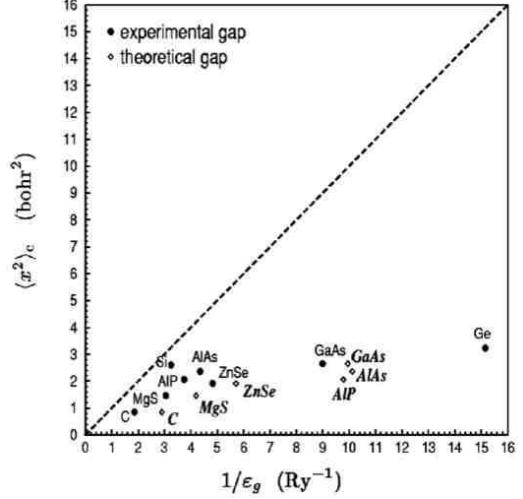


FIG. 5: Diagonal element of the KS localization tensor vs. the inverse direct gap (theoretical and experimental), for several elemental and binary semiconductors (from Ref. [44]). The points corresponding to Si and Ge with the theoretical gaps are out of scale. From Ref. [44].

Correlated (Mott) insulators

Starting from the noninteracting Hamiltonian of Eq. (74) and augmenting it with an on-site repulsive term we get the two-band Hubbard model

$$\hat{H} = \sum_{j\sigma} [(-1)^j \Delta c_{j\sigma}^\dagger c_{j\sigma} - t(c_{j\sigma}^\dagger c_{j+1\sigma} + \text{H.c.})] + U \sum_j n_{j\uparrow} n_{j\downarrow}. \quad (78)$$

The explicitly correlated ground-state wavefunction has been found by exact diagonalization [13], and the corresponding $\langle x^2 \rangle_c$ has been computed as a function of U for fixed $t/\Delta = 1.75$. The results are shown in Fig. 6 in dimensionless units; it turns out that there is only one singular point $U = 2.27t$, where $\langle x^2 \rangle_c$ diverges. Indeed, it has been verified that at such value the ground-state becomes degenerate with the first excited singlet state, i.e. the system is metallic. The singular point is the fingerprint of a quantum phase transition: on the left we have a band-like insulator, and on the right a Mott-like insulator. The two insulating states are *qualitatively* different; by adopting the modern jargon, nowadays we could say that they are *topologically* distinct. The static ionic charges (on anion and cation) are continuous across the transition, while the dynamical (Born) effective charge on a given site changes sign [68]. Other studies of the localization tensor within the same Hubbard model can be found in Ref. [69].

The transition from a band metal to a Mott insulator has been studied in a model linear chain of Li atoms by Vetere *et al.* [49]. At a mean-field level the infinite chain is obviously metallic at any lattice constant a , since there is one valence electron per cell. However the mean-

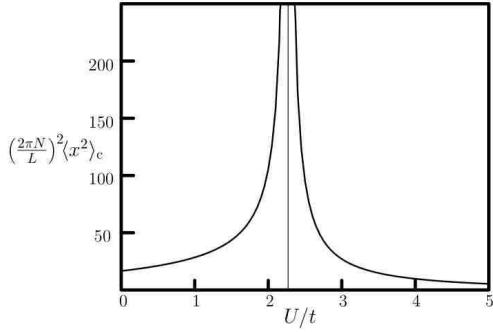


FIG. 6: Squared localization length for the Hamiltonian in Eq. (78) at half filling for $t/\Delta = 1.75$. The system undergoes a quantum phase transition from band-like insulator to Mott-like insulator at $U/t = 2.27$. From Ref. [13]

field description becomes inadequate at large a , where the electrons localize and the system becomes a Mott insulator. If electron correlation is properly accounted for at any a , the system undergoes a sharp metal-insulator transition at a critical a .

The calculations addressed linear Li_N systems (N up to 8), where the finite size prevents a sharp transition; the tradeoff is that full configuration interaction was affordable with 6 atomic orbitals per site (yielding more than 10^9 symmetry-adapted Slater determinants). The wavefunction of Vetrone *et al.* is therefore exempt from any bias insofar as the treatment of correlation is concerned, although its quality is determined by the basis set. A study of the longitudinal component $\langle x^2 \rangle_c$ of the localization tensor indicates rather clearly the occurrence of the metal-insulator transition at $a \simeq 7$ bohr; other indicators give concordant results [49]. For comparison, the nearest-neighbour distance in 3d metallic lithium is 5.73 bohr.

Disordered (Anderson) insulators

We start from the same Hamiltonian as in Eq. (74), and we replace the ordered string $(-1)^j$ by a random string of ± 1 , chosen with equal (and uncorrelated) probability. This system models a random binary alloy at 50% concentration. It is well known both from analytical arguments and actual simulations that its spectrum is gapless [6, 70]. The density of states for both the ordered and disordered systems are shown in Fig. 7, and confirm the expected features. The band structure of Eq. (75) yields obviously a gapped density of states; at the band edges it shows van Hove singularities, which in 1d have the character of $1/\sqrt{\epsilon}$ divergences. As discussed above, the system is insulating at half filling and conducting otherwise. The disordered system, instead, is gapless and nonetheless insulating at any filling. In fact, this model Hamiltonian describes a paradigmatic Anderson

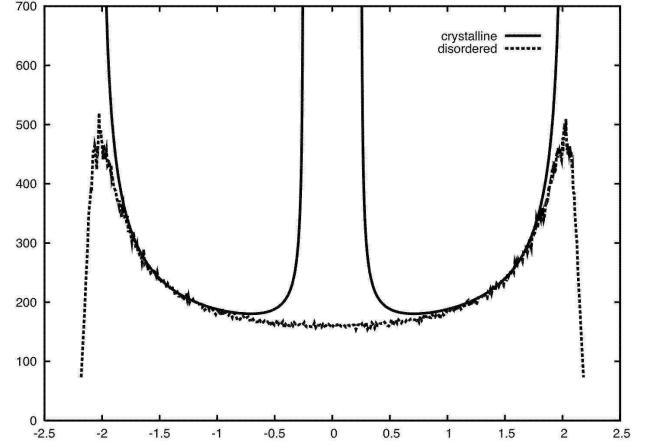


FIG. 7: Density of states (arbitrary units) for a model binary alloy in 1d. The crystalline (band) case corresponds to the Hamiltonian of Eq. (74) with $\Delta = 0.25$ and $t = 1$. The disordered (Anderson) case corresponds to a random choice of the anion/cation distribution.

insulator in 1d.

The conventional theory of transport focusses on the nature of the one-particle orbitals at the Fermi level; in Anderson insulators these are localized, thus forbidding steady state currents [4]. More than fifty years of literature have been devoted to investigate Anderson insulators under the most diverse aspects. [6, 70–72].

At variance with such wisdom, a recent work has addressed this paradigmatic Anderson insulator from the nonconventional viewpoint of the modern theory of the insulating state [50]. In the spirit of Kohn's theory the individual Hamiltonian eigenstates become apparently irrelevant, while the focus is on the many-electron ground state as a whole. The squared localization length $\langle x^2 \rangle_c$ has been computed within OBCs from Eq. (76), and found to be finite, as expected. Nonetheless its value is about 20 times larger than the one for the band insulator, at the same value of the parameters (i.e. $\Delta = 0.25$, $t = 1$). This reflects the fact that the scattering mechanisms are profoundly different: incoherent (Anderson) versus coherent (band). In the latter case, the Hamiltonian eigenstates are individually conducting but “locked” by the Pauli principle if the Fermi level lies in the gap.

Quantum Hall insulators

The results of a typical quantum Hall (QH) experiment are shown in Fig. 8. The most perspicuous feature is that whenever the 2d electron fluid is in the QH regime: (i) the transverse resistivity is quantized; and (ii) the longitudinal resistivity vanishes. While the plots are about the integer QH effect, the same two features are common to the fractional QH effect as well.

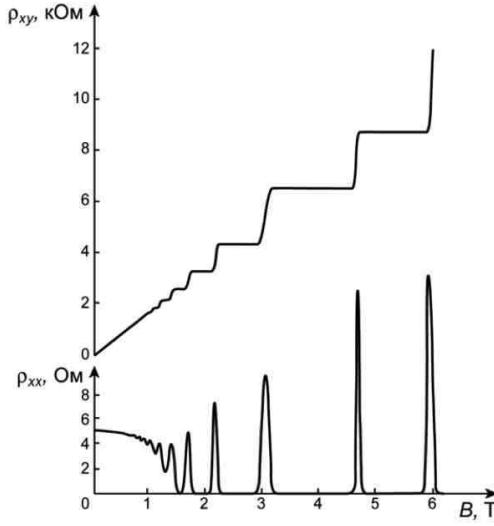


FIG. 8: The results of a typical quantum Hall experiment.

We switch from resistivity to conductivity; in isotropic 2d-systems the relationship is

$$\rho_{11} = \frac{\sigma_{11}}{\sigma_{11}^2 + \sigma_{12}^2} \quad \rho_{12} = -\frac{\sigma_{12}}{\sigma_{11}^2 + \sigma_{12}^2}, \quad (79)$$

hence in the QH regime $\sigma_{11} = 0$ and $\sigma_{12} = -1/\rho_{12}$ is quantized. Since the static longitudinal conductivity vanishes, the system is by definition insulating: this is why we speak of “QH insulators”.

We are going to show that electron localization—as defined in the modern theory of the insulating state—is the common *cause* for both the vanishing of the dc longitudinal conductivity and the quantization of the transverse one. Therefore to predict whether the dc transverse conductivity is quantized, it is enough to inspect electron localization in the ground state. This outstanding result was proved in 2005 [37].

Our system is a square of size $L \times L$, where the presence of a macroscopic B field *forbids* simple PBCs; this, however, is no serious problem. We choose the Landau gauge for the magnetic vector potential in Eq. (29), and we impose the magnetic boundary conditions which are customary in the QH literature [73]; this requires the total flux BL^2 to be an integer times Φ_0 , where $\Phi_0 = hc/e$ is the flux quantum.

When switching from 3d to 2d, many of the above formulae require obvious modifications. For an isotropic system, the 2d analogue of Eqs. (58) and (59) can be written as

$$\int_0^\infty \frac{d\omega}{\omega} \text{Re } \sigma_{11}(\omega) = \frac{\pi e^2 N}{\hbar L^2} \text{Re } \langle x^2 \rangle_c, \quad (80)$$

$$\text{Re } \sigma_{12}(0) = \frac{2e^2 N}{\hbar L^2} \text{Im } \langle xy \rangle_c. \quad (81)$$

A glance at the sum-over-states expression for $\langle r_\alpha r_\beta \rangle_c$, Eq. (45), shows that whenever $\text{Re } \langle r_\alpha r_\beta \rangle_c$ converges,

then $\text{Im } \langle r_\alpha r_\beta \rangle_c$ converges as well. However, we are going to show that the latter may only converge to quantized values in the present 2d case.

By definition we have

$$\text{Im } \langle xy \rangle_c = \frac{1}{N} \text{Im } \langle \partial_1 \Psi_0(0) | \partial_2 \Psi_0(0) \rangle. \quad (82)$$

In the large- L limit, we may write

$$\begin{aligned} \text{Im } & \langle \partial_1 \Psi_0(0) | \partial_2 \Psi_0(0) \rangle \\ &= \left(\frac{L}{2\pi} \right)^2 \text{Im} \int_0^{\frac{2\pi}{L}} d\kappa_1 \int_0^{\frac{2\pi}{L}} d\kappa_2 \langle \partial_1 \Psi_0(\kappa) | \partial_2 \Psi_0(\kappa) \rangle. \end{aligned} \quad (83)$$

The dimensionless integral is over a closed surface (a torus), given the magnetic boundary conditions. It is therefore equal to $-\pi C_1$, where C_1 is the Chern number of the first class (Sect.); hence

$$\text{Im } \langle xy \rangle_c = -\frac{L^2}{4\pi N} C_1, \quad (84)$$

and the transverse conductivity, Eq. (81) becomes

$$\text{Re } \sigma_{12}(0) = -\frac{e^2}{2\pi\hbar} C_1 = -\frac{e^2}{h} C_1. \quad (85)$$

We have thus arrived at the famous Niu-Thouless-Wu formula [73] which holds for both the integer and fractional QH effect. The original derivation was based on an analysis of the Green function, under the hypothesis that the system has a Fermi gap; in the present derivation the presence of a Fermi gap becomes apparently irrelevant, since the localization tensor is a pure *ground-state* property. Underlying the modern theory of the insulating state, however, is a fluctuation-dissipation theorem, relating the ground state to the excitations of the system.

The main message of Ref. [37] (and of the present Section) can be summarized by saying that *any* 2d insulator displays a quantized transverse conductance (nonvanishing only in absence of time-reversal symmetry).

An electron fluid is kept in the QH regime by disorder, and analytical implementations of the present formulae are obviously not possible. Nonetheless, in order to illustrate how the theory works, it is expedient to consider the academic case of noninteracting electrons in a flat substrate potential. If the first Landau level is fully occupied (i.e. at filling $\nu = 1$) the electron density is uniform and equal to $n_0 = 1/(2\pi\ell^2)$, where $\ell = (\hbar c/eB)^{1/2}$ is the magnetic length; the modulus of the density matrix has the simple expression (for single occupancy)

$$|\rho(\mathbf{r}, \mathbf{r}')| = |P(\mathbf{r}, \mathbf{r}')| = n_0 e^{-|\mathbf{r}-\mathbf{r}'|^2/(4\ell^2)}, \quad (86)$$

and the cumulant second moment is clearly finite. The trace $\langle x^2 \rangle_c + \langle y^2 \rangle_c$ of the localization tensor is in fact

equal to ℓ^2 , the squared magnetic length. If the filling is fractional the density cannot be uniform (for noninteracting electrons), and, more important, the cumulant second moment diverges (see Ref. [37] for details): the system is not insulating, and its transverse conductivity is not quantized.

Chern insulators

The QH state of matter, discovered in 1980 and addressed in the previous Section, provides the first example of a quantum state which is topologically distinct from all previously known states of matter. This is the reason why its macroscopic quantization properties are very robust (“topologically protected”), and insensitive to small changes in material parameters. The debut of geometrical concepts in electronic structure theory dates from two milestone papers: TKNN in 1982 [24] and Niu-Thouless-Wu in 1985 [73]. These papers addressed a 2d electron fluid (noninteracting and interacting, respectively) in the presence of a macroscopic magnetic field: the Hamiltonian, therefore, cannot be lattice periodical.

A subsequent breakthrough on the theory side is the Haldane model Hamiltonian [61]: this can be considered as the precursor of modern topological insulators [31]. Its trademark is quantized Hall conductance in absence of a macroscopic magnetic field.

The model is comprised of a 2d honeycomb lattice with two tight-binding sites per primitive cell with site energies $\pm\Delta$, real first-neighbor hopping t_1 , and complex second-neighbor hopping $t_2 e^{\pm i\phi}$, as shown in Fig. 9. Within this two-band model, one deals with insulators by taking the lowest band as occupied. The appeal of the model is that the vector potential and the Hamiltonian are lattice periodical and the single-particle orbitals always have the usual Bloch form. Essentially, the microscopic magnetic field can be thought as staggered (i.e. up and down in different regions of the cell), but its cell average vanishes.

For a 2d lattice-periodical Hamiltonian the Chern number C_1 has been defined before, Eq. (73). As a function of the flux parameter ϕ , this system undergoes a transition from a normal insulator ($C_1 = 0$) to Chern insulator ($|C_1| = 1$).

Like in the QH case discussed in Sect. : (i) the imaginary part of the localization tensor is related to the Chern number by Eq. (84); and (ii) the trace of the localization tensor is finite. This is confirmed by the simulations of Ref. [51], where the actual value of Ω_I , Eq. (71), is computed. We remind that, despite Ω_I being finite, localized Wannier functions (with finite quadratic spread) do not exist in Chern insulators.

The Haldane model insulator and the QH insulator both display quantized transverse conductivity, and both

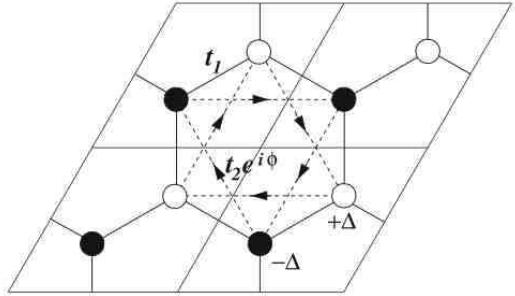


FIG. 9: Four unit cells of the Haldane model [61]. Filled (open) circles denote sites with $E_0 = -\Delta$ ($+\Delta$). Solid lines connecting nearest neighbors indicate a real hopping amplitude t_1 ; dashed arrows pointing to a second-neighbor site indicates a complex hopping amplitude $t_2 e^{i\phi}$. Arrows indicate sign of the phase ϕ for second-neighbor hopping.

are localized in the sense of the modern theory of the insulating state. It is worth pointing out, though, that the decay of the density matrix is qualitatively different: exponential in the Haldane case [51] vs. Gaussian in the (noninteracting) QH case, Eq. (86).

No microscopic realization of a Chern insulator (in absence of a macroscopic magnetic field) is known to date. A *mesoscopic* Chern insulator, in the same spirit as the Haldane model, was synthesized in 2008 [74]. Chern insulators remained a curiosity of academic interest only for many years. In 2005 it was realized that the Haldane model is the forebear of a completely new class of insulators, called “topological insulators” [25, 31].

Topological insulators

Although Chern insulators do not exist in nature, a different kind of topological insulators *does exist* in nature in both 2d and 3d [26–31]; their macroscopic quantization properties are topologically protected, and insensitive to small changes in material parameters, exactly like in the QH case.

Modern topological insulators are characterized by novel invariants called \mathbb{Z}_2 (close relatives of the Chern number in 2d and of the Chern invariant in 3d). The paradigmatic \mathbb{Z}_2 topological insulator is the Kane-Mele model Hamiltonian in 2d [25]. Most important, at variance with Chern insulators, the modern topological insulators are time-reversal symmetric.

Since spin-orbit interaction plays a major role, the definition of localization tensor needs to be augmented to explicitly include the spin coordinates; this seems to be rather straightforward, and is deferred to further publications. After this is done, the scope of the modern theory of the insulating state would include topological insulators.

Obviously, no investigation of this kind exists yet.

Nonetheless, a closely related investigation concerning the Wannier functions in a \mathbb{Z}_2 insulator is appearing these days [52]. The results confirm that the ground state of a topological insulator is localized in the sense of the modern theory of the insulating state: Ω_I , Eq. (71), is in fact finite. Furthermore—at variance with the Chern-insulator case—localized Wannier functions (with finite quadratic spread) do exist, although a generalization of the Marzari-Vanderbilt theory [52, 60] is needed.

CONCLUSIONS

According to the milestone 1964 Kohn’s paper [7], the organization of the electrons in their *ground state* is qualitatively different in insulators and metals. This paper provided the foundations of the “theory of the insulating state”. The modern developments of the theory started many years later (1999) with the work of Resta and Sorella [13], and continue to these days.

Here we provide a comprehensive presentation of the modern theory. In so doing, we emphasize the geometrical properties of the many-body wavefunction, which were somewhat hidden in most of the original literature.

We focus on the linear response of the many-body wavefunction to an infinitesimal “twist” (or “flux”) in the electronic Hamiltonian. This is the ground-state property which discriminates insulators from metals. The distance between the (infinitesimally) twisted and untwisted ground states, measured by the quantum metric tensor in the thermodynamic limit, is finite in insulators and divergent in metals.

Technically, in our formulation the twist has the dimensions of an inverse length; the quantum metric tensor has therefore the dimensions of a squared length and is extensive. The key intensive quantity is the metric tensor *per electron*; this is generally called “localization tensor” or “second cumulant moment” of the electron distribution. It is a global property of the electronic ground state, apparently unrelated to the system excitations. In some cases (discussed above) the localization tensor is naturally endowed with an imaginary antisymmetric part, whose physical meaning is also outstanding.

The modern theory of the insulating state encompasses all kinds of known insulators. Here we discuss, using uniform concepts and notations, the most diverse cases: open and periodic boundary conditions; interacting and noninteracting electrons, crystalline and disordered systems; time-reversal symmetry present and absent. We even briefly comment on topological insulators.

Work partly supported by the ONR grant N00014-07-1-1095.

- [1] F. Bloch, Z. Phys. **52**, 555 (1928).
- [2] A. H. Wilson, Proc. Roy. Soc. A **133**, 458, and **134**, 277 (1931).
- [3] N. F. Mott, Proc. Phys. Soc. (London) **62**, 416 (1949).
- [4] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
- [5] N. Mott, *Metal-Insulator Transitions*, 2nd ed. (Taylor & Francis, London, 1990).
- [6] E. Abrahams (Ed.), *50 Years of Anderson Localization*, (World Scientific, Singapore, 2010).
- [7] W. Kohn, Phys. Rev. **133**, A171 (1964).
- [8] W. Kohn, in *Many-Body Physics*, edited by C. DeWitt and R. Balian (Gordon and Breach, New York, 1968), p. 351.
- [9] Only about 80 citations (ISI) in the 25 years from 1964 to 1991.
- [10] D. Vanderbilt and R. Resta, in: *Conceptual foundations of materials: A standard model for ground- and excited-state properties*, S.G. Louie and M.L. Cohen, eds. (Elsevier, 2006), p. 139.
- [11] R. Resta and D. Vanderbilt, in: *Physics of Ferroelectrics: a Modern Perspective*, Topics in Applied Physics Vol. **105**, Ch. H. Ahn, K. M. Rabe, and J.-M. Triscone, eds. (Springer-Verlag, 2007), p. 31.
- [12] R. Resta, J. Phys.: Condens. Matter **22** 123201 (2010).
- [13] R. Resta and S. Sorella, Phys. Rev. Lett. **82**, 370 (1999).
- [14] I. Souza, T. Wilkens, and R. M. Martin, Phys. Rev. B **62**, 1666 (2000).
- [15] R. Resta, J. Phys.: Condens. Matter **14**, R625 (2002).
- [16] D. Bures, Trans. Am. Math. Soc. **135**, 199 (1969).
- [17] M. V. Berry, Proc. Roy. Soc. Lond. A **392**, 45 (1984).
- [18] *Geometric Phases in Physics*, edited by A. Shapere and F. Wilczek (World Scientific, Singapore, 1989).
- [19] D. J. Thouless, *Topological Quantum Numbers in Nonrelativistic Physics* (World Scientific, Singapore, 1998).
- [20] A. Bohm, A. Mostafazadeh, H. Koizumi, Q. Niu, and J. Zwanziger, *The Geometric Phase in Quantum Systems* (Springer, Berlin, 2003).
- [21] R. Resta, J. Phys.: Condens. Matter **12**, R107 (2000).
- [22] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. **82**, 1959 (2010).
- [23] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, Reading, 1994), p.140.
- [24] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
- [25] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
- [26] D. N. Sheng, Z. Y. Weng, L. Sheng, and F. D. M. Haldane, Phys. Rev. Lett. **97**, 036808 (2006).
- [27] S.-C. Zhang, Physics **1**, 6 (2008).
- [28] Y. L. Chen et al., Science **325**, 178 (2009).
- [29] J. E. Moore, Physics **2**, 82 (2009).
- [30] X. L. Qi and S. C. Zhang, Phys. Today **63**(1), 38 (2010).
- [31] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
- [32] J. P. Provost and G. Vallee, Commun. Math. Phys. **76**, 289 (1980).
- [33] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959), reprinted in Ref. [18], p.104.
- [34] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures in Physics*, Vol. 2 (Addison Wesley,

Reading, 1964), Sect. 15-4.

[35] R. Resta, Phys. Rev. Lett. **80**, 1800 (1998).

[36] R. Resta, J. Chem. Phys. **124**, 104104 (2006).

[37] R. Resta, Phys. Rev. Lett. **95**, 196805 (2005).

[38] D. N. Zubarev, Soviet Phys. Ushpekhi **3**, 320 (1960).

[39] D. N. Zubarev, *Non-Equilibrium Statistical Mechanics* (Consultants Bureau, New York, 1974).

[40] R. McWeeny, *Methods of Molecular Quantum Mechanics*, Second Edition (Academic, London, 1992).

[41] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II, Nonequilibrium Statistical Mechanics*, Springer Series in Solid-State Sciences, Vol. **31**, (Springer, Berlin, 1985).

[42] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin, Reading, 1975).

[43] R. Resta, Phys. Rev. Lett. **96**, 137601 (2006).

[44] C. Sgarirovello, M. Peressi, and R. Resta, Phys. Rev. **64**, 115202 (2001).

[45] M. Veithen, X. Gonze, and Ph. Ghosez, Phys. Rev. B **66**, 235113 (2002).

[46] N. D. M. Hine and W. M. C. Foulkes, J. Phys: Condens. Matter **19**, 506212 (2007).

[47] A. Monari, G. L. Bendazzoli, and S. Evangelisti, J. Chem. Phys. **129**, 134104 (2008).

[48] C. Aebischer, D. Baeriswyl, and R. M. Noack , Phys. Rev. Lett. **86**, 468 (2001).

[49] V. Vetere, A. Monari, G.L. Bendazzoli, S. Evangelisti, and B. Paulus, J. Chem. Phys. **128**, 214701 (2008).

[50] G. L. Bendazzoli, S. Evangelisti, A. Monari, and R. Resta, J. Chem. Phys. **133**, 064703 (2010).

[51] T. Thonhauser and D. Vanderbilt, Phys. Rev. B **74**, 235111 (2006).

[52] A. A. Soluyanov and D. Vanderbilt, Phys. Rev. B, in press.

[53] E. Akkermans, J. Math. Phys. **38**, 1781 (1997).

[54] W. Kohn, Phys. Rev. Lett. **2**, 393 (1959).

[55] J. des Cloizeaux, Phys. Rev. **135**, A685 (1964); *ibid.* **135**, A697 (1964).

[56] S. Ismail-Beigi and T.A. Arias, Phys. Rev. Lett. **82**, 2127 (1999).

[57] L. He and D. Vanderbilt, Phys. Rev. Lett. **86**, 5341 (2001).

[58] G. F. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, 2005).

[59] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).

[60] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[61] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).

[62] D. Ceresoli, T. Thonhauser, D. Vanderbilt, and R. Resta, Phys. Rev. B **74**, 024408 (2006).

[63] S. Coh and D. Vanderbilt, Phys. Rev. Lett. **102**, 107603 (2009).

[64] D. J. Thouless, J. Phys. C **17**, L325 (1984).

[65] C. Brouder, G. Panati, M. Calandra, Ch. Mourougane, and N. Marzari, Phys. Rev. Lett. **98**, 046402 (2007).

[66] J. G. Àngyà, Int. J. Quantum Chem. **109**, 2340 (2009).

[67] J. G. Àngyà, Curr. Org. Chem., in press.

[68] R. Resta and S. Sorella, Phys. Rev. Lett. **87**, 4738 (1995).

[69] T. Wilkens and R. M. Martin, Phys. Rev. B **63**, 235108 (2001).

[70] B. Kramer and A. MacKinnon, Rep. Prog. Phys. **56**, 1469 (1993).

[71] D. J. Thouless, Phys. Rep. **13**, 93 (1974).

[72] A. Lagendijk, B. van Tiggelen, and D. S. Wiersma, Phys. Today **62**(8), 24 (2009).

[73] Q. Niu, D. J. Thouless, and Y. S. Wu, Phys. Rev. B **31**, 3372 (1985).

[74] M. Taillefumier, V. K. Dugaev, B. Canals, C. Lacroix, and P. Bruno, Phys. Ref. B **78**, 155330 (2008).